Least Squares Problems in Orthornormalization

Shanwen Hu*
Research Center for Operator Algebras
East China Normal University, Shanghai 200241, P.R. China

Abstract

For any n-tuple $(\alpha_1, \cdots, \alpha_n)$ of linearly independent vectors in Hilbert space H, we construct a unique orthonormal basis $(\epsilon_1, \cdots, \epsilon_n)$ of $span\{\alpha_1, \cdots, \alpha_n\}$ satisfying:

$$\sum_{i=1}^{n} \|\epsilon_i - \alpha_i\|^2 \leq \sum_{i=1}^{n} \|\beta_i - \alpha_i\|^2$$

for all orthonormal basis $(\beta_1, \cdots, \beta_n)$ of $span\{\alpha_1, \cdots, \alpha_n\}$. We study the stability of the orthornormalization and give some applications and examples.

Key words: orthornormalization, Gram-Schmidt orthogonalization, least square

2010 *Mathematics Subject Classification*: 46C05,15A60,65J05

1 Notations and Introduction

Throughout this paper,

1. $\mathbb{C}$ (or $\mathbb{R}$) is the complex (or real) number field.

2. For any $z \in \mathbb{C}$, $\overline{z}$ is the complex conjugate of z. $Re(z)$ is the real part of z.

3. $M_{n,\,m}(\mathbb{C}) = \{(a_{ij})\}$ is the set of $n \times m$ complex matrices. $M_n(\mathbb{C}) = M_{n,\,n}(\mathbb{C})$.

4. $\mathbb{C}^n = M_{n,1}(\mathbb{C})$. The identity of $\mathbb{C}^n$ denoted by I_n.

5. For any $(a_{ij}) \in M_{n,\,m}(\mathbb{C})$, $\overline{(a_{ij})} = (\overline{a}_{ij}) \in M_{n,\,m}(\mathbb{C})$.

6. The standard basis of $\mathbb{C}^n$ is denoted by $\mathbf{e}_i = (\delta_{1i}, \cdots, \delta_{in}), i = 1, \cdots, n$.

7. For any $(a_{ij}) \in M_n(\mathbb{C})$. $tr(a_{ij}) = \sum_{i=1}^{n} a_{ii}$. $\sigma\big((a_{ij})\big)$ is the set of spectrum of (a_{ij}). $\|(a_{ij})\| = \sup\{\|(a_{ij})x^T\| : x \in \mathbb{C}^n, \|x\| = 1\}$.

8. H is a complex Hilbert space with the inner product $\langle \cdot, \cdot \rangle$.

9. $H^n = \{(\alpha_1, \cdots, \alpha_n) : \alpha_i \in H, i = 1, \cdots, n\}$ is a Hilbert space, the inner product defined by: $\langle (\alpha_1, \cdots, \alpha_n), (\beta_1, \cdots, \beta_n) \rangle = \sum_{i=1}^{n} \langle \alpha_i, \beta_i \rangle$.

10. For any $(a_{ij}) \in M_{n,\,m}(\mathbb{C}), (a_{ij})^T = (a_{ji}) \in M_{m,\,n}(\mathbb{C})$ and for any $(\alpha_1, \cdots, \alpha_n) \in H^n$,

$$(\alpha_1, \cdots, \alpha_n)^T = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

*E-mail: swhu@math.ecnu.edu.cn

If $\{\alpha_1, \cdots, \alpha_n\} \subset H^n$ is subset of linearly independent vectors in H. The orthonormalization of $\{\alpha_1, \cdots, \alpha_n\}$ is to find a solution $\{\beta_1, \cdots, \beta_n\}$ in $span\{\alpha_1, \cdots, \alpha_n\}$ to the system: for any i, j,

$$\langle \beta_i, \beta_j \rangle = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}, \quad i, j = 1, \cdots, n.$$

The orthonormalization can be carried out in infinitely ways. The most simple and practical tool is the Gram-Schmidt process. It is a recursive process and are used widely in various fields. However in general, the Gram-Schmidt process can not provide a simple construction formula for the solution and can not give a method to calculate the sum of squares

$$\|(\beta_1, \cdots, \beta_n) - (\alpha_1, \cdots, \alpha_n)\|^2 = \sum_{i=1}^{n} \|\beta_i - \alpha_i\|^2.$$

Gram-Schmidt process is also unstable due to the repeated various operations. These restrict its applications, especially in the abstract or theoretical analysis.

In numerical linear algebra, Householder method is also used in the orthornormalization. The Gram-Schmidt process produces the jth orthogonalized vector after the jth iteration, while Householder method produces all vectors only at the end. And theoretically Household method take twice operations as Gram-Schmidt process, but it uses orthogonal transformation at each iteration, so it is stable. The Household method is restricted in numerical linear algebra only.

The Gram-Schmidt process and Householder method can be find in Linear Algebra and Matrix Analysis text books, for example, see [1] or [2].

In Section 2, we provide a simple and uniform formula $K(\alpha_1, \cdots, \alpha_n)$ (see (2.4)) for any Hilbert space, on complex field or real field, with finite or infinite dimensional, in numerical form or not in numerical form, no iteration, to construct an orthonormal basis of $span\{\alpha_1, \cdots, \alpha_n\}$, only according the direct information $\{\langle \alpha_i, \alpha_j \rangle : i, j = 1, \cdots, n\}$, satisfying:

$$\sum_{i=1}^{n} \|\epsilon_i - \alpha_i\|^2 = n + \sum_{i=1}^{n} \|\alpha_i\|^2 - 2tr\big((\langle \alpha_i, \alpha_j \rangle)^{1/2}\big).$$

Moreover, we show $n + \sum_{i=1}^{n} \|\alpha_i\|^2 - 2tr\big((\langle \alpha_i, \alpha_j \rangle)^{1/2}\big)$ is the lower bound of all sum squares of orthonormal basis of $span\{\alpha_1, \cdots, \alpha_n\}$, and $(\epsilon_1, \cdots, \epsilon_n)$ is the unique one minimized the sum of squares.

In Section 3, we show our construction is stable in the sense: for any given $\epsilon > 0$ and any linearly independent n-tuple $(\alpha_1, \cdots, \alpha_n)$ in H^n, there exists δ, dependent only on $(\alpha_1, \cdots, \alpha_n)$ and ϵ, such that for any $(\beta_1, \cdots, \beta_n) \in H^n$, $\max\{\|\alpha_i - \beta_i\|\} < \delta$ implies $\|K(\alpha_1, \cdots, \alpha_n) - K(\beta_1, \cdots, \beta_n)\| < \epsilon$.

In section 4, as an application, we establish a formula for the distance of between any $\gamma \in H$ and $span\{\alpha_1, \cdots, \alpha_n\}$, generalize the one in the case $\{\alpha_1, \cdots, \alpha_n\}$ is an orthonormal basis. For any ϵ-mutually orthonormal basis $(\alpha_1, \cdots, \alpha_n)$,(see (4.4) for the definition), we show there is an orthonormal basis $(\epsilon_1, \cdots, \epsilon_n)$, the distance between them in H^n not exceed $\sqrt{2(n-1)\epsilon}$.

2 Least Square in Orthornormalization

(1) For any $(\alpha_1, \cdots, \alpha_n) \in H^n$, $(a_1, \cdots, a_n) \in \mathbb{C}^n$, if $\eta = \sum_{i=1}^n a_i \alpha_i$, we write

$$\eta = (a_1, \cdots, a_n)(\alpha_1, \cdots, \alpha_n)^T = (\alpha_1, \cdots, \alpha_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

(2) If for any i, $\beta_i = (a_{i1}, \cdots, a_{in})(\alpha_1, \cdots, \alpha_n)^T, i = 1, \cdots, n$, then we write

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}. \tag{2.1}$$

That is

$$\beta_i = \mathbf{e}_i (a_{ij})(\alpha_1, \cdots, \alpha_n)^T, i = 1, \cdots, n.$$

It is easy to check: if $\eta = \sum_i^n a_i \alpha_i$ and $\zeta = \sum_{j=1}^m b_j \beta_j$, then

$$\langle \eta, \zeta \rangle = \sum_{j=1}^m \sum_{i=1}^n a_i \bar{b}_j \langle \alpha_j, \beta_i \rangle = (\bar{b}_1, \cdots, \bar{b}_m)(\langle \alpha_j, \beta_i \rangle)(a_1, \cdots, a_n)^T. \tag{2.2}$$

Theorem 2.1. *Suppose $(\alpha_1, \cdots, \alpha_n)$ is n-tuple of linearly independent vectors in a complex Hilbert space, then*

(1) For any orthonormal base $(\beta_1, \cdots, \beta_n)$ of span$\{\alpha_1, \cdots, \alpha_n\}$, there exists an invertible $(a_{ij}) \in M_n(\mathbb{C})$ such that (2.1) holds and satisfying,

$$\sum_{i=1}^n \|\beta_i - \alpha_i\|^2 = n + \sum_{i=1}^n \|\alpha_i\|^2 - 2Re\Big(tr\big((\langle \alpha_j, \alpha_i \rangle)(a_{ji})\big)\Big). \tag{2.3}$$

(2) $(\langle \alpha_j, \alpha_i \rangle)$ is positive definite and if $K(\alpha_1, \cdots, \alpha_n) = (\epsilon_1, \cdots, \epsilon_n)$ defined by

$$\begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} = \begin{pmatrix} \langle \alpha_1, \alpha_1 \rangle & \cdots & \langle \alpha_1, \alpha_n \rangle \\ \vdots & & \vdots \\ \langle \alpha_n, \alpha_1 \rangle & \cdots & \langle \alpha_n, \alpha_n \rangle \end{pmatrix}^{-1/2} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \tag{2.4}$$

then

(a) $(\epsilon_1, \cdots, \epsilon_n)$ is a orthonormal basis of span$\{\alpha_1, \cdots, \alpha_n\}$.

(b)

$$\sum_{i=1}^n \|\epsilon_i - \alpha_i\|^2 = n + \sum_{i=1}^n \|\alpha_i\|^2 - 2tr\Big(\big(\langle \alpha_j, \alpha_i \rangle\big)^{1/2}\Big). \tag{2.5}$$

(c) For any orthonormal base $(\beta_1, \cdots, \beta_n)$ of span$\{\alpha_1, \cdots, \alpha_n\}$,

$$\sum_{i=1}^n \|\epsilon_i - \alpha_i\|^2 \le \sum_{i=1}^n \|\beta_i - \alpha_i\|^2,$$

and the equality holds if and only if $(\beta_1, \cdots, \beta_n) = (\epsilon_1, \cdots, \epsilon_n)$.

Proof. (1) There is no problem for the existence and the invertibility of (a_{ij}). Noticing $\beta_i = \mathbf{e}_i(a_{ij})(\alpha_1, \cdots, \alpha_n)^T$ and $\alpha_i = \mathbf{e}_i(\alpha_1, \cdots, \alpha_n)^T$, applying (2.2), we have

$$\begin{aligned}
&\sum_{i=1}^{n} \|\beta_i - \alpha_i\|^2 \\
= \;& n + \sum_{i=1}^{n} \|\alpha_i\|^2 - \sum_{i=1}^{n} \Big(\langle \beta_i, \alpha_i \rangle + \langle \alpha_i, \beta_i \rangle \Big) \\
= \;& n + \sum_{i=1}^{n} \|\alpha_i\|^2 - 2Re\Big(\sum_{i=1}^{n} \big\langle \mathbf{e}_i(a_{ij})(\alpha_1, \cdots, \alpha_n)^T, \mathbf{e}_i(\alpha_1, \cdots, \alpha_n)^T \big\rangle \Big) \\
= \;& n + \sum_{i=1}^{n} \|\alpha_i\|^2 - 2Re\Big(\sum_{i=1}^{n} \mathbf{e}_i \big((\langle \alpha_j, \alpha_i \rangle)(a_{ij})^T \big) \mathbf{e}_i^T \Big) \\
= \;& n + \sum_{i=1}^{n} \|\alpha_i\|^2 - 2Re\Big(tr\big((\langle \alpha_j, \alpha_i \rangle)(a_{ji}) \big) \Big).
\end{aligned}$$

(2) (a) For any $(a_1, \cdots, a_n) \in \mathbb{C}^n$, not all zero, by (2.1)

$$(\overline{a}_1, \cdots, \overline{a}_n)(\langle \alpha_j, \alpha_i \rangle)(a_1, \cdots, a_n)^T = \Big\langle \sum_{i=1}^{n} a_i \alpha_i, \sum_{i=1}^{n} a_i \alpha_i \Big\rangle > 0,$$

so $(\alpha_j, \alpha_i\rangle)$ is positive definite. Noticing $\overline{(\langle \alpha_i, \alpha_j \rangle)^{-1/2}} = (\overline{\langle \alpha_i, \alpha_j \rangle})^{-1/2}$, and applying (2.1), we have

$$\langle \epsilon_i, \epsilon_j \rangle = \mathbf{e}_j (\overline{\langle \alpha_i, \alpha_j \rangle})^{-1/2} (\langle \alpha_j, \alpha_i \rangle) \big((\langle \alpha_i, \alpha_j \rangle)^T \big)^{-1/2} \mathbf{e}_i^T = \mathbf{e}_j \mathbf{e}_i^T = \delta_{ji},$$

for any i, j.

(b) (2.5) follows from (2.3) and $(a_{ji}) = (\langle \alpha_j, \alpha_i \rangle)^{-1/2}$.

(c) There exists a unitary $U = (u_{ij}) \in M_n(\mathbb{C})$ such that

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}.$$

Then

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} \langle \alpha_1, \alpha_1 \rangle & \cdots & \langle \alpha_1, \alpha_n \rangle \\ \vdots & & \vdots \\ \langle \alpha_n, \alpha_1 \rangle & \cdots & \langle \alpha_n, \alpha_n \rangle \end{pmatrix}^{-1/2} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

so we have

$$(a_{ij}) = (u_{ij}) \big(\langle \alpha_i, \alpha_j \rangle \big)^{-1/2} \qquad or \qquad (a_{ji}) = (\langle \alpha_j, \alpha_i \rangle)^{-1/2} (u_{ji}).$$

Therefore

$$Re\Big(tr\big((\langle \alpha_j, \alpha_i \rangle)(a_{ji}) \big) \Big) = Re\Big(tr\big((\langle \alpha_j, \alpha_i \rangle)^{1/2} (u_{ji}) \big) \Big).$$

To the end of the proof, it is enough to show the statement: for any positive $T \in M_n(\mathbb{C})$ and any unitary U, $Re\big(tr(TU)\big) \leq tr(T)$, and the equation holds if and only if $U = I_n$. Since T may be assumed diagonal, it is easy to prove the statement. We omit the details. ∎

Remark 2.2. *If H is a real complete inner product space, the conclusions in the Theorem 2.1 are still true. In fact, the inverse and squares root of a symmetric matrix are all real. Noticing $\langle \alpha_i, \alpha_j \rangle = \langle \alpha_j, \alpha_i \rangle$, the proof has nothing to be changed except the unitary matrix replaced by a orthogonal matrix.*

Example 2.3. *Suppose $H = \mathbb{R}^m$, $\alpha_i = (a_{i1}, \cdots, a_{im}), i = 1, \cdots, n$. If $(\alpha_1, \cdots, \alpha_n)$ is linearly independent, then $n \le m$ and*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

has rank n, so AA^T is invertible. If $(AA^T)^{-1/2}A = (h_{ij})$, then $K(\alpha_1, \cdots, \alpha_n) = (\epsilon_1, \cdots, \epsilon_n)$ defined by: $\epsilon_i = (h_{i1}, \cdots, h_{im}), i = 1, \cdots, n$.

For any other orthonormal basis $(\beta_1, \cdots, \beta_n)$ of span$\{\alpha_1, \cdots, \alpha_n\}$, if

$$\beta_i = (b_{i1}, \cdots, b_{im}), i = 1, \cdots, n,$$

then

$$\sum_{i=1}^{n} \|\epsilon_i - \alpha_i\|^2 = \sum_{j=1}^{m}\sum_{i=1}^{n} |a_{ij} - h_{ij}|^2 \quad < \quad \sum_{j=1}^{m}\sum_{i=1}^{n} |b_{ij} - h_{ij}|^2 = \sum_{i=1}^{n} \|\beta_i - \alpha_i\|^2.$$

Example 2.4. *Let $H = L^2([0,1])$ be the space of real function which square is Lebesgue integrable. For any n, $\{1, x, \cdots, x^n\}$ is linearly independent. By the definition of (2.4),*

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_n \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n-1} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \cdots & \cdots \\ \frac{1}{3} & \cdots & \cdots & \cdots & \cdots \\ \vdots & & & & \\ \frac{1}{n-1} & \cdots & \cdots & \cdots & \frac{1}{2n-1} \end{pmatrix}^{-1/2} \begin{pmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{n-1} \end{pmatrix}$$

If the transformation matrix denoted by A_n, then

$$\sum_{i=1}^{n} \|\alpha_i - \epsilon_i\|^2 = n + 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} - 2tr(A_n^{-1}).$$

For $n = 4$, with the help of MATLAB, we immediately get:

$$\begin{cases} \epsilon_1 = 1.8145 - 2.8273x + 2.0557x^2 - 0.6986x^3 \\ \epsilon_2 = -2.8273 + 18.1940x - 26.7823x^2 + 11.9872x^3 \\ \epsilon_3 = 2.0557 - 26.7823x + 64.5308x^2 - 39.9282x^3 \\ \epsilon_4 = -0.6986 + 11.9872x - 39.9282x^2 + 32.5816x^3 \end{cases}$$

and

$$\sum_{i=1}^{4} \|x^{i-1} - \epsilon_i\|^2 = 4 + 1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} - 2tr(A_4^{-1}) = 2.2201,$$

the accuracy is controlled by MATLAB.

3 Perturbations

In this section, we will discuss the perturbation problem and show our construction is stable in some sense.

Note although $(\langle \alpha_i, \alpha_j \rangle)$ is non singular, it may be nearly singular when $\|\alpha_i\|$ is too small. We may modify our work, replacing $(\alpha_1, \cdots, \alpha_n)$ by $(\frac{\alpha_1}{\|\alpha_1\|}, \cdots, \frac{\alpha_n}{\|\alpha_n\|})$ in this section.

To show the main result of the section, we need the following Lemmas. Lemma 3.1 is well known.

Lemma 3.1. *If A is an invertible element of unital C^*-algebra $\mathcal{A}$, for any $B \in \mathcal{A}$ with $\|B - A\| < \|A^{-1}\|^{-1}$, then B is invertible satisfying*

$$\|A^{-1} - B^{-1}\| \leq \frac{\|A^{-1}\|^2 \|A - B\|}{1 - \|A^{-1}\| \|A - B\|} \tag{3.1}$$

and

$$\|B^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|A - B\|}. \tag{3.2}$$

Lemma 3.2. *If A, B are positive and invertible on a Hilbert space H, then*

$$\|A^{-1/2} - B^{-1/2}\| \leq \|A^{1/2}\| \|A^{-1} - B^{-1}\|. \tag{3.3}$$

Proof. It is enough to show

$$\|A - B\| \leq \|A^{-1}\| \|A^2 - B^2\|. \tag{3.4}$$

There exists $\lambda \in \sigma(A - B)$ with $|\lambda| = \|A - B\|$ and $\{x_n\} \subset H, \|x_n\| = 1, i = 1, 2, \cdots$ such that

$$\lim_{n \to \infty} \|(A - B)x_n - \lambda x_n\| = 0.$$

Since $A^2 - B^2 = A(A - B) + (A - B)A - (A - B)(A - B)$, we have

$$\begin{aligned}
& \|A^2 - B^2\| \\
\geq\ & \liminf_{n \to \infty} \left| \left\langle (A^2 - B^2))x_n, x_n \right\rangle \right| \\
=\ & \liminf_{n \to \infty} \left| \left\langle A(A - B)x_n, x_n \right\rangle + \left\langle (A - B)Ax_n, x_n \right\rangle - \left\langle (A - B)(A - B)x_n, x_n \right\rangle \right| \\
=\ & \liminf_{n \to \infty} \left| \left\langle \lambda A x_n, x_n \right\rangle + \left\langle \lambda A x_n, x_n \right\rangle - \left\langle \lambda (A - B)x_n, x_n \right\rangle \right| \\
=\ & \liminf_{n \to \infty} \left| |\lambda| \left| \left\langle A x_n, x_n \right\rangle + \left\langle B x_n, x_n \right\rangle \right| \right| \\
\geq\ & \|A - B\| \|A^{-1}\|^{-1},
\end{aligned}$$

thus we obtain (3.4) and (3.3). □

The special conditions of the following Lemma 3.3 are prepared for the proof of Theorem 3.4.

Lemma 3.3. *Suppose $A, B, C \in M_n(\mathbb{C})$, B is positive, A is positive and invertible with $\|A\| \leq n, 1 \leq \|A^{-1}\|$. If there exists η with $0 \leq \eta \leq 2\|A^{-1}\|^{-1})$, $\|A - B\| < \eta$ and $\|A - C\| < \eta$, then B is invertible and*

$$\|A^{-1/2} C B^{-1/2} - I_n\| \leq 2n(n + 1)\|A^{-1}\|^2 \eta. \tag{3.5}$$

Proof. It follows Lemma 3.1, B is invertible and

$$\|A^{-1} - B^{-1}\| < 2\|A^{-1}\|^2\|A - B\| \le 2\|A^{-1}\|^2\eta, \tag{3.6}$$

$$\|B^{-1}\| \le 2\|A^{-1}\|. \tag{3.7}$$

Condition $\|A^{-1}\| \ge 1$ implies $\|A^{-1}\| \ge \|A^{-1/2}\|$ and

$$\|B^{-1/2}\| = \|B^{-1}\|^{1/2} \le (2\|A^{-1}\|)^{1/2} \le 2\|A^{-1}\|. \tag{3.8}$$

By (3.3) and (3.6),

$$\|A^{-1/2} - B^{-1/2}\| \le \|A^{1/2}\|\|A^{-1} - B^{-1}\| \le 2\|A^{1/2}\|\|A^{-1}\|^2\eta.$$

At last, applying (3.7) and (3.8), we get

$$\begin{aligned} & \|A^{-1/2}CB^{-1/2} - I_n\| \\ \le\ & \|A^{-1/2}CB^{-1/2} - A^{-1/2}AB^{-1/2}\| + \|A^{1/2}B^{-1/2} - A^{1/2}A^{-1/2}\| \\ \le\ & \|A^{-1/2}\|\|C - A\|\|B^{-1/2}\| + \|A^{1/2}\|\|B^{-1/2} - A^{-1/2}\| \\ \le\ & 2\|A^{-1}\|^2\eta + 2\|A\|\|A^{-1}\|^2\eta \\ =\ & 2(n+1)\|A^{-1}\|^2\eta. \end{aligned}$$

□

Theorem 3.4. *Suppose $(\alpha_1, \cdots, \alpha_n)$ is an n-tuple of linearly independent units of H^n, for any given $\epsilon > 0$, let*

$$\delta = \big(8n^2(n+1)\|(\langle\alpha_j, \alpha_i\rangle)^{-1}\|^2\big)^{-1}\epsilon,$$

then for any $(\beta_1, \cdots, \beta_n)$ of units in H^n, if

$$\max\{\|\alpha_i - \beta_i\| : i, j = 1, \cdots, n\} < \delta, \tag{3.9}$$

then $(\beta_1, \cdots, \beta_n)$ is linearly independent and satisfying

$$\|K(\alpha_1, \cdots, \alpha_n) - K(\beta_1, \cdots, \beta_n)\|^2 < \epsilon. \tag{3.10}$$

Proof. Let $A = (\langle\alpha_j, \alpha_i\rangle), B = (\langle\beta_j, \alpha_i\rangle), C = (\langle\beta_j, \beta_i\rangle)$. Since $tr(A) = n$, so $\lambda_{min} \le 1$, where $\lambda_{min} = \min\big(\sigma(A)\big)$, this implies $\|A^{-1}\| = \lambda_{min}^{-1} \ge 1$. $\|A\| \le n$ follows from for any i, j, $|\langle\alpha_i, \alpha_j\rangle| \le 1$. The condition (3.9) guarantee for any i, j,

$$|\langle\alpha_j, \alpha_i\rangle - \langle\beta_j, \beta_i\rangle\| < 2\delta, \quad |\langle\alpha_j, \alpha_i\rangle - \langle\alpha_j, \beta_i\rangle| < \delta.$$

Consequently,

$$\|A - B\| < 2n\delta < \big(2\|A^{-1}\|^2\big)^{-1}, \qquad \|A - C\| < 2n\delta.$$

Let $\eta = 2n\delta$, then all conditions in Lemma 3.3 are all satisfied, so we have

$$\|A^{-1/1}CB^{-1/2} - I_n\| \le 2(n+1)\|A^{-1}\|^2\eta = 4n(n+1)\|A^{-1}\|^2\delta.$$

The linearly independence of $(\beta_1, \cdots, \beta_n)$ follows from the invertibility of B.

Now suppose $K(\alpha_1, \cdots, \alpha_n) = (\epsilon_1, \cdots, \epsilon_n), K(\beta_1, \cdots, \beta_n) = (\tau_1, \cdots, \tau_n)$, then by the definition (2.4) and formula (2.2),

$$
\begin{aligned}
& \sum_{i=1}^{n} \|\epsilon_i - \tau_i\|^2 = 2n - \sum_{i=1}^{n} \big(\langle \epsilon_i, \tau_i \rangle + \langle \tau_i, \epsilon_i \rangle\big) \\
= \quad & 2n - \sum_{i=1}^{n} 2Re\Big(\mathbf{e}_i(\langle \alpha_j, \alpha_i \rangle)^{-1/2}(\langle \beta_j, \alpha_i \rangle)(\langle \beta_j, \beta_i \rangle)^{-1/2}\mathbf{e}_i^T\Big) \\
= \quad & 2Re\Big(tr(I - A^{-1/2}CB^{-1/2})\Big) \\
\leq \quad & 2\|I_n - A^{-1/2}CB^{-1/2}\| tr(I_n) \leq 8n^2(n+1)\|A^{-1}\|^2\delta = \epsilon,
\end{aligned}
$$

thus we complete the proof. □

4 Applications

In this section, we will give some applications of our construction (2.4) to the theoretical analysis.

For a fixed linearly independent vectors $\{\alpha_1, \cdots, \alpha_n\}$, for any $\gamma \in H$, we define

$$
D(\gamma) = dist(\gamma, span\{\alpha_1, \cdots, \alpha_n\}) = \inf\{\|\gamma - \beta\| : \beta \in span\{\alpha_1, \cdots, \alpha_n\}\}.
$$

If $\{\alpha_1, \cdots, \alpha_n\}$ is mutual orthogonal units, then

$$
D(\gamma) = \sqrt{\|\gamma\|^2 - \sum_{i=1}^{n} |\langle \gamma, \alpha_i \rangle|^2}. \tag{4.1}
$$

The following theorem will show (4.1) is just (4.2) in the special case.

For simple, in this section, for any $(\alpha_1, \cdots, \alpha_n) \in H^n$, we define

$$
(\alpha_1, \cdots, \alpha_n) \circ (\alpha_1, \cdots, \alpha_n) = \big(\langle \alpha_i, \alpha_j \rangle\big) \in M_n(\mathbb{C}).
$$

Theorem 4.1. *Suppose $(\alpha_1, \cdots, \alpha_n)$ is n-tuple linearly independent vectors in a complex Hilbert space, then*

$$
D(\gamma) = \sqrt{\frac{det\big((\gamma, \alpha_1 \cdots, \alpha_n) \circ (\gamma, \alpha_1, \cdots, \alpha_n)\big)}{det\big((\alpha_1, \cdots, \alpha_n) \circ (\alpha_1, \cdots, \alpha_n)\big)}}. \tag{4.2}
$$

Proof. Suppose $(\epsilon_1, \cdots, \epsilon_n) = K(\alpha_1, \cdots, \alpha_n)$ defined by (2.4). Let $\Delta = det(\langle \alpha_i, \alpha_j \rangle)$, applying formula (2.2) and $\sum_{i=1}^{n} \mathbf{e}^T \mathbf{e}_i = I_n$, we have

$$
\begin{aligned}
& \sum_{i=1}^{n} |\langle \gamma, \epsilon_i \rangle|^2 \\
= \quad & \sum_{i=1}^{n} |\mathbf{e}_i(\langle \alpha_j, \alpha_i \rangle)^{-1/2}(\langle \gamma, \alpha_1 \rangle, \cdots, \langle \gamma, \alpha_n \rangle)^T|^2 \\
= \quad & \sum_{i=1}^{n} (\langle \alpha_1, \gamma \rangle, \cdots, \langle \alpha_n, \gamma \rangle)(\langle \alpha_j, \alpha_i \rangle)^{-1/2}\mathbf{e}_i^T \mathbf{e}_i(\langle \alpha_j, \alpha_i \rangle)^{-1/2}(\langle \gamma, \alpha_1 \rangle, \cdots, \langle \gamma, \alpha_n \rangle)^T \\
= \quad & \big(\langle \alpha_1, \gamma \rangle, \cdots, \langle \alpha_n, \gamma \rangle)(\langle \alpha_j, \alpha_i \rangle)^{-1}(\langle \gamma, \alpha_1 \rangle \cdots \langle \gamma, \alpha_n \rangle\big)^T. \qquad (*)
\end{aligned}
$$

Let A_{ij} be the (i, j) cofactors in $(\langle \alpha_j, \alpha_i \rangle)$ (not in $(\langle \alpha_i, \alpha_j \rangle)$!), then

$$
\begin{aligned}
(*) &= \frac{1}{\Delta}(\langle\alpha_1,\gamma\rangle,\cdots,\langle\alpha_n,\gamma\rangle)\begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn}\end{pmatrix}\begin{pmatrix}\langle\gamma,\alpha_1\rangle \\ \vdots \\ \langle\gamma,\alpha_n\rangle\end{pmatrix} \\
&= \frac{1}{\Delta}\Big(\sum_{j=1}^{n}\sum_{i=1}^{n} A_{ji}\langle\gamma,\alpha_i\rangle\langle\alpha_j,\gamma\rangle\Big) \\
&= \frac{-1}{\Delta} det\begin{pmatrix} 0 & \langle\gamma,\alpha_1\rangle & \cdots & \langle\gamma\,\alpha_1\rangle \\ \langle\alpha_1,\gamma\rangle & \langle\alpha_1,\alpha_1\rangle & \cdots & \langle\alpha_1,\alpha_n\rangle \\ \vdots & & & \vdots \\ \langle\alpha_n,\gamma\rangle & \langle\alpha_n,\alpha_1\rangle & \cdots & \langle\alpha_n,\alpha_n\rangle\end{pmatrix} \\
&= \|\gamma\|^2 - \frac{det((\gamma,\alpha_1,\cdots,\alpha_n)\circ(\gamma,\alpha_1,\cdots,\alpha_n))}{\Delta}.
\end{aligned}
$$

Since

$$dist(\gamma, span\{\alpha_1,\cdots,\alpha_n\}) = \sqrt{\|\gamma\|^2 - \sum_{i=1}^{n}|\langle\gamma,\epsilon_i\rangle|^2},$$

we obtain (4.2). ☐

Corollary 4.2. *Suppose $\{\alpha_1,\alpha_2,\cdots\}$ is a sequence of independent vectors, then for any $\gamma\in H$, the distance between γ and the closure of $span\{\alpha_1,\alpha_2,\cdots\}$ is:*

$$\lim_{n\to\infty}\sqrt{\frac{det((\gamma,\alpha_1\cdots,\alpha_n)\circ(\gamma,\alpha_1,\cdots,\alpha_n))}{det((\alpha_1,\cdots,\alpha_n)\circ(\alpha_1,\cdots,\alpha_n))}}.$$

Lemma 4.3. *Suppose $E=\{\lambda_i, i=1,\cdots,n\}$ is a finite positive numbers set with $\sum_{i=1}^{n}\lambda_i = n$ and $\max\{|1-\lambda_i| : i=1,\cdots,n\} = \epsilon \le \frac{1}{2(n-1)}$, then*

$$\Big|\sum_{i=1}^{n}\lambda_i^{1/2} - n\Big| \le \epsilon. \tag{4.3}$$

Proof. We may assume $\lambda_1 = \min\big(E\big)$ and $\lambda_n = \max\big(E\big)$.

If $\lambda_1 = 1$ or $\lambda_n = 1$, then for all $i, \lambda_i = 1$, there is nothing to do. If $\lambda_n = 1+\eta$, then $0<\eta\le\epsilon$. Since the function

$$\sum_{i=1}^{n-2}\lambda_i^{1/2} + \big(n-1-\eta-\sum_{i=1}^{n-2}\lambda_i\big)^{1/2}$$

on

$$[1-\frac{1}{2(n-1)}, 1+\frac{1}{2(n-1)}]\times\cdots\times[1-\frac{1}{2(n-1)}, 1+\frac{1}{2(n-1)}]\subset\mathbb{R}^{n-2}$$

obtains its maximum value only when $\lambda_i = 1-\frac{\eta}{n-1}, i=1,\cdots,n-2$, consequently,

$$\sum_{i=1}^{n}\lambda_i^{1/2} \le (1+\eta)^{1/2} + \sum_{i=1}^{n-1}\Big(1-\frac{\eta}{n-1}\Big)^{1/2} \le n+\eta.$$

If $\lambda_1 = 1 - \zeta$, similar argument show

$$\sum_{i=1}^{n} \lambda_i^{1/2} \geq n - \zeta.$$

Then (4.3) follows the assumption $\eta \leq \epsilon$ and $\zeta \leq \epsilon$.

$\square$

If $\{\alpha_1, \cdots, \alpha_n\}$ satisfies condition,

$$\epsilon = \max\{|\langle \alpha_i, \alpha_j \rangle| : i, j = 1, \cdots, n\}, \tag{4.4}$$

we will say $\{\alpha_1, \cdots, \alpha_n\}$ is ϵ-mutually orthogonal.

In [3], Hu and Xue proved, if $\{\alpha_1, \cdots, \alpha_n\}$ is ϵ-mutually orthogonal, then there are mutually orthogonal $\{\beta_1, \cdots, \beta_n\}$ with $\|\alpha_i - \beta_i\| < 6(n-1)\epsilon, i = 1, \cdots, n$. Now, we have the following:

Theorem 4.4. *Suppose n-tuple units $(\alpha_1, \cdots, \alpha_n)$ in H is ϵ-mutually orthogonal with $\epsilon < \frac{1}{2(n-1)}$, then there is an orthonormal basis $(\epsilon_1, \cdots, \epsilon_n)$ of span$\{\alpha_1, \cdots, \alpha_n\}$ such that*

$$\sum_{i=1}^{n} \|\epsilon_i - \alpha_i\|^2 < 2(n-1)\epsilon. \tag{4.5}$$

Proof. Let $T = (\langle \alpha_i, \alpha_j \rangle)$ and assume $\sigma(T) = \{\lambda_1, \cdots, \lambda_n\}$. Since $\langle \alpha_i, \alpha_i \rangle = 1$ and

$$\sum_{j \neq i} |\langle \alpha_i, \alpha_j \rangle| < (n-1)\epsilon,$$

by Gersgorin Theorem ([2]Theorem 6.1.1), for all $i, |\lambda_i - 1| \leq (n-1)\epsilon$. Meanwhile,

$$\sum_{i=1}^{n} \lambda_i = tr(T) = \sum_{i=1}^{n} \langle \alpha_i, \alpha_i \rangle = n,$$

then by the Lemma 4.1,

$$\Big| tr(T^{1/2}) - n \Big| = \Big| \sum_{i=1}^{n} \lambda_i^{1/2} - n \Big| \leq (n-1)\epsilon.$$

Let $(\epsilon_1, \cdots, \epsilon_n) = K(\alpha_1, \cdots, \alpha_n)$ defined by (2.4), applying (2.5)

$$\sum_{i=1}^{n} \|\epsilon_i - \alpha_i\|^2 = 2\big| tr(T^{1/2}) - n \big| = 2\Big| \sum_{i=1}^{n} \lambda_i^{1/2} - n \Big| \leq 2(n-1)\epsilon.$$

$\square$

Acknowledgements The author would like to thank Professor Huaxin Lin for his helpful comments and suggestions from their multiple discussions.

References

[1] Leon,S.J, Linear Algebra with Applications, Sixth Edition(ISBN 0-13-033781-1), Person Education Asia Limited and China Machine Press, 2004.

[2] Roger A.Horn and Charles R.Johnson, Matrix Analysis, Cambridge University Press

[3] S.W.Hu and Y.F.Xue, Completeness of n-tuple of Projections in C^*-algebras, arXiv:1210.4670